

Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/EP05/003366

International filing date: 24 March 2005 (24.03.2005)

Document type: Certified copy of priority document

Document details: Country/Office: EP
Number: 04090431.0
Filing date: 12 November 2004 (12.11.2004)

Date of receipt at the International Bureau: 25 May 2005 (25.05.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)



World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse

24. 03. 2005



Europäisches
Patentamt

European
Patent Office

Office européen
des brevets

Bescheinigung

Certificate

Attestation

Die angehefteten Unterla-
gen stimmen mit der
ursprünglich eingereichten
Fassung der auf dem näch-
sten Blatt bezeichneten
europäischen Patentanmel-
dung überein.

The attached documents
are exact copies of the
European patent application
described on the following
page, as originally filed.

Les documents fixés à
cette attestation sont
conformes à la version
initialement déposée de
la demande de brevet
européen spécifiée à la
page suivante.

Patentanmeldung Nr. Patent application No. Demande de brevet n°

04090431.0

Der Präsident des Europäischen Patentamts;
Im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets
p.o.

R C van Dijk



Anmeldung Nr:
Application no.: 04090431.0
Demande no:

Anmeldetag:
Date of filing: 12.11.04
Date de dépôt:

Anmelder/Applicant(s)/Demandeur(s):

Epigenomics AG
Kleine Präsidentenstrasse 1
10178 Berlin
ALLEMAGNE

Bezeichnung der Erfindung/Title of the invention/Titre de l'invention:
(Falls die Bezeichnung der Erfindung nicht angegeben ist, siehe Beschreibung.
If no title is shown please refer to the description.
Si aucun titre n'est indiqué se référer à la description.)

Rapid analysis of CpG methylation patterns using RNase T1 cleavage and MALDI-TOF

In Anspruch genommene Priorität(en) / Priority(ies) claimed /Priorité(s)
revendiquée(s)
Staat/Tag/Aktenzeichen/State/Date/File no./Pays/Date/Numéro de dépôt:

Internationale Patentklassifikation/International Patent Classification/
Classification internationale des brevets:

C12Q1/68

Am Anmeldetag benannte Vertragstaaten/Contracting states designated at date of
filing/Etats contractants désignées lors du dépôt:

AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IS IT LU MC
NL PL PT RO SE SI SK TR LI

Rapid analysis of CpG methylation patterns using RNase T1 cleavage and MALDI-TOF

Philipp Schatz, Dimo Dietrich and Matthias Schuster*

The authors wish to be known that, in their opinion, the first two authors should be regarded as joint
First Authors

* To whom correspondence should be addressed. Tel: +49 30 24345100; Fax: +49 30 24345299;
Email: schuster@epigenomics.com

Epigenomics AG, Science Department, Kleine Präsidentenstraße 1, D-10178 Berlin, Germany

Here we introduce a method for the fast and accurate analysis of DNA methylation based on bisulfite treated DNA. The target region is PCR amplified using a T7 RNA polymerase promoter tagged primer. A subsequent *in vitro* transcription leads to a transcript which contains guanosine residues only at sites that contained methylated cytosines before bisulfite treatment. In a single tube reaction using guanosine-specific cleavage by RNase T1, a specific pattern of RNA fragments is formed. This pattern directly represents the methylation state of the sample DNA and is analyzed using MALDI-TOF technology. This method was successfully applied to the analysis of artificially methylated and unmethylated DNA, mixtures thereof and colon DNA samples. The applicability for the analysis of both PCR products and cloned PCR products is demonstrated. The observed methylation patterns were confirmed by bisulfite sequencing.

Introduction

DNA methylation in mammals is almost exclusively observed in position 5 of the cytosine base of the dinucleotide CpG. These dinucleotides are underrepresented in

the genome and appear at only 20% of the expected frequency, as they are selected against during evolution due to their high mutability. Nevertheless, the genome contains CpG islands with a CpG dinucleotide frequency much higher than statistically expected. Promoter CpG methylation directly influences the transcriptional activity of the respective gene, e.g. by interfering with the binding of transcription factors (1) and is involved in fundamental processes like X-chromosome inactivation (2), imprinting (3), tumorigenesis (4), aging and cell differentiation (5). Efficient methods are needed to analyze this epigenetic information. Most techniques for the analysis of methylation patterns depend on a preceding bisulfite treatment of the template DNA leading to a deamination of all unmethylated cytosines to uracil, leaving only methylated cytosines unaltered (6). Thus the bisulfite reaction transforms the methylation information into sequence information which is easily analyzed using standard molecular biology techniques, most notably PCR. A powerful method for analyzing multiple CpG sites is bisulfite sequencing, which can be performed either directly on PCR products or on cloned PCR products generated on bisulfite treated DNA. As direct bisulfite PCR sequencing reflects the mean of all subpopulations within a given sample bisulfite sequencing of cloned PCR products is the method of choice for accurate analysis of the methylation patterns present in a heterogeneous sample. It is also the method of choice for detecting stretches of comethylated CpG positions which often carry biological information when occurring in promoter regions. Since many clones have to be sequenced in parallel and since only a fraction of the sequence trace information, namely the polymorphisms at CpG positions, are relevant for methylation analysis, a method which targets this information specifically and which can be run in high throughput at low cost, is highly desirable. The method presented in this paper is based on *in vitro* transcription of bisulfite PCR products, base-specific cleavage and final MALDI-TOF analysis of the obtained RNA fragments. It represents a novel approach towards the efficient and low cost analysis of methylation pattern and is amenable to high throughput analysis. Fragmented RNA is a viable analyte for MALDI analysis, as it is more easily analyzed than DNA of the same size (7). The RNA fragment pattern can be correlated to the methylation pattern of the template DNA. In previous studies the combination of the above mentioned methods has already been successfully applied to comparative sequence analysis (8), discovery of single nucleotide polymorphisms (SNPs, 9, 10) and to the characterization of short tandem repeats (STR, 11).

Material and Methods

DNA Preparation

Methylated DNA for the experiment described in Figure 2 was prepared by treating human genomic DNA (Promega) with SssI methyltransferase (New England Biolabs) in the presence of S-adenosyl-methionine according to the manufacturer's instructions. Unmethylated DNA was prepared by MDA (multiple displacement amplification), a genome-wide amplification method described by Dean *et al.* (12). For the preparation of mixtures with defined methylation states a portion of this amplificate was treated with SssI as described above and mixed with the unmethylated amplificate to give mixtures with 0, 20, 40, 50, 60, 80 and 100% methylation in all CpG positions. DNA from colon tissue samples (Biocat) was extracted using the QIAamp DNA mini kit protocol (Qiagen). Bisulfite treatment was performed as previously described (13).

PCR Amplification, Cloning and Sequencing

A 289 bp fragment within the promoter region of the CDH13 gene, encoding H-cadherin was PCR amplified with the primers 5'-tcttttcTTTGTATTAGGTTGGAAGTGGT-3' (forward) and 5'-gtaatacgactcactatagggagCCCAAATAAATCAACAACAACA-3' (reverse). Gene-specific sequences are marked in capital letters and functional tags are given in small letters.

PCR was performed in a total volume of 25 µl containing 1 U Hotstar Taq polymerase (Qiagen), 12,5 pmol of forward and reverse primers, 1x PCR buffer (Qiagen), 0,2 mM of each dNTP (Fermentas) and 5ng template. PCR of clones was performed using picked colony cells without DNA purification. Cycling was performed using a Mastercycler (Eppendorf) under the following conditions: 15 min at 95°C and 40 cycles at 95°C for 1 min, 55°C for 45 s and 72°C for 1 min.

The PCR product of the CDH13 fragment of SssI methylated and bisulfite treated DNA was purified using the QIAquick kit (Qiagen) and subsequently cloned into the pGEM[®]-T Vector System (Promega) following the manufacturer's recommendations.

Plasmid DNA from one clone (pEPI2383) was extracted using the Plasmid Mini Kit (Qiagen) protocol.

PCR products and plasmid pEPI2383 (restricted as described below) were sequenced using BigDye chemistry (Applied Biosystems) according to the manufacturer's recommendations.

RNA Transcription and Base-specific Cleavage

Prior to T7 transcription the plasmid DNA from clone pEPI2383 was hydrolyzed in 10 μ l volume using 10 U Eco52I (Fermentas) in the recommended buffer. The reaction was performed for 1 h at 37°C. Inactivation was performed for 10 min at 65°C.

RNA transcription was performed in a 25 μ l volume containing 10 μ l PCR product or hydrolyzed plasmid DNA, respectively, containing 20 U T7 RNA polymerase (Fermentas), 1x transcription buffer (Fermentas), 0,5 mM of each NTP (Fermentas) and 4 U RNase inhibitor (Fermentas). After an incubation of 1,5 h at 37°C 4 U of RNase T1 (Fermentas) was added immediately and incubated for another 45 min at 37°C.

MALDI-TOF Analysis

Prior to the MALDI-TOF analysis the RNA was desalted by adding 25 mg Clean Resin (Sequenom) to each preparation and incubating for 20 min at room temperature. The cleavage reaction mixture was diluted 5 fold and 0,5 μ l was mixed with 0,5 μ l of organic matrix (saturated 3-hydroxypicolinic acid in 50% acetonitrile containing 0,05 M dibasic ammonium citrate) on a ScoutTM384 stainless steel target plate. Fragment analysis was carried out on a Biflex III mass spectrometer (Bruker Daltonics). Mass spectra were recorded in the negative ion reflector time-of-flight mode. Potentials were 19 kV for the linear and 20 kV for the reflector acceleration. IS/2 potential was set at 17,75 kV. Ion extraction was delayed by 300 ns. The detector was gated to prevent saturation by molecules smaller than 700 Da. Usually 60 shots were accumulated per sample spot, smoothed using a Golay-Savitzky filter and baseline-corrected. The interpretation of fragment masses was done using the software 'Mongo Oligo Mass Calculator v2.05' (<http://medlib.med.utah.edu/masspec/mongo.htm>).

Results

Methodological Background

The presented novel approach for the detection of DNA methylation patterns requires bisulfite treatment of the template DNA. During the bisulfite treatment only unmethylated cytosines are converted into uracils (Fig. 1, I). Thus, the methylation information is translated into sequence information. The double stranded template DNA is converted into two single stranded molecules the cytosines of which represent the methylation information in the template DNA. In the subsequent PCR (Fig. 1, II) which is performed using a reverse primer containing a T7 RNA polymerase promoter tag (gtaatacgactcactatagggag), a complementary strand with low guanosine content is generated. The guanosines within this strand resemble the methylation information of the template DNA. After T7-mediated transcription (Fig. 1, III) these guanosine residues are subject to cleavage by RNase T1 from *Aspergillus oryzae*. Thus, the RNA fragmentation pattern corresponds to the methylation pattern of the original DNA (Fig. 1, IV). This methylation fingerprint is visualized by MALDI-TOF mass spectrometry. In addition, the T7 transcript is marked by a control tag at the 3'-end introduced by the forward primer. Using this control tag the successful full length transcription and a successful RNase T1 cleavage are easily monitored. For validation of this method a 289 bp fragment of the promoter region of the CDH13 gene was analyzed. This fragment contained 13 cytosines within a CpG context, all of which are methylated after SssI treatment.

Sequencing

PCR products derived from methylated and unmethylated bisulfite treated DNA and the plasmid pEPI2383 were sequenced. In addition, the *EcoR*52I hydrolyzed plasmid pEPI2383 was sequenced. The expected sequences of the corresponding RNA *in vitro* transcripts are depicted in Figure 2, I., A-D.

RNA Fragmentation

Figure 2 (II.) shows the m/z values of all fragments expected after T1 cleavage of transcript A (Fig. 2, I., A) generated from completely methylated template DNA. In addition to the gene-specific fragments 1-14 the RNase T1 cleavage should result in the fragments C1, C2 and P1 - P4. With the exception of those labeled *italic* all

fragments indeed were detected by MALDI-TOF analysis (Fig. 2, III., A). Fragments remaining undetected were either too small and thus not distinguishable from matrix background noise or too large and outside the detection range of the instrument.

In contrast to transcript A the T7 transcript derived from the originally unmethylated DNA contains no guanosines within the gene-specific sequence (Fig. 2, I., B). Thus, the subsequent G-specific cleavage generated only one gene-specific fragment ($m/z = 90913$) which is not detectable due to its large size. The control fragment C1 ($m/z = 1991$) was successfully detected with MALDI -TOF (Fig. 2, III., B), demonstrating successful transcription and fragmentation. Again, fragments C2 and P1 - P4 were not detectable because of their sizes.

The MALDI-TOF spectrum of the fragmented transcript derived from a PCR product of the plasmid pEPI2383 is shown in Figure 2 (III., C). According to the virtual sequence of this transcript (Fig. 2, I., C), the spectrum differs in some additional and lacking fragments when compared to the spectrum derived from methylated DNA (Fig. 2, III., A). Fragment 13, caused by methylation of CpG sites at position 258 and 269 is lacking, which is easily explained by the additional guanosine at position 266 (Fig. 2, I., C). Additional guanosines are expected at low frequency due to incomplete bisulfite conversion. Thus fragment 13 is cleaved into two fragments. Of these $m/z = 2603$ is detected whereas $m/z = 980$ was not detectable due to its small size.

Fragment 12, generated by cleavage at positions 210 and 258 also possessed an internal cleaving site (position 242). Thus two additional fragments of $m/z = 5166$ and $m/z = 10103$ appeared instead of fragment 12 ($m/z = 15253$).

The fragment pattern of the digested transcript derived from the PCR product of pEPI2383 is devoid of fragment 1 due to the unmethylated CpG site at position 32 (Fig. 2, I., C) which is explained by incomplete methylation of the template DNA. Hence, a fragment was generated which reflected the sum of fragments 1 and 2 ($m/z = 24327$). This fragment was too large to be detectable.

The MALDI-TOF spectra of the fragmented T7 transcript derived from pEPI2383 without prior PCR amplification is shown in Figure 2 (III., D). As expected this spectrum corresponds well to the spectrum derived from the PCR product of this plasmid with the exception of an additional peak R which is observed due to the additional sequence (position 300-311 in Fig. 2, I.) between the multiple cloning site of the vector and the recognition site of the applied restriction enzyme *EcoR52I*. This sequence attached to the end of the transcript resulted in 3 additional fragments, of

which only fragment R (Fig. 2, III., D) was detected, since the other two ($m/z = 651$ and $m/z = 243$) are outside the detection range.

Analysis of colon DNA samples

To demonstrate the viability of the novel method for the analysis of clinical sample material over a wide range of methylation levels colon DNA samples were analyzed. Abberant methylation of the CDH13 promoter has been implicated with colorectal cancer (14). For the purpose of this study two tumor DNA samples showing high methylation levels and two normal colon DNA samples showed low methylation levels were selected. Ten clones each of the amplified promoter region of the CDH13 gene were analyzed in direct comparison to sequencing (Figure 3).

There is excellent correlation of the methylation states of those CpG positions which are analyzed unambiguously by both methods for predominantly methylated (T1, T2) as well as for predominantly unmethylated samples (N1, N2). In some cases sequencing failed to reveal the methylation state in position 32, 258 and 269 which are either close to the sequencing primer or to the end of the sequence. On the other hand the mass range limitation of the MALDI instrument used in this study did not enable the unambiguous assignment of the methylation state of all CpG positions, since the absence of a fragment cannot be interpreted as the absence of methylation in a particular position unless this is justified by detection of larger fragments. In clones A and J of sample T1, e.g., the presence of fragment 6+7 ($m/z=5117$, Supplemental Material) is caused by methylation in positions 122 and 138 surrounding the unmethylated position 136. When several neighboring CpG positions are unmethylated the resulting fragments become larger and are not easily detected. Positions 154 and 210 are considered to be non-analyzable, because the corresponding fragments are either too large to be detected reliably or too short to be distinguished from the background. However, this is not a general limitation. State of the art instrumentation has been described to detect RNAs as long as 2180 nucleotides (15) and has been applied to sequence 50-mer to 100-mer DNA and RNA fragments (16).

Finally an aliquot of the same bisulfite treated colon DNA samples was analyzed directly, i.e. without prior cloning. For comparison mixtures of unmethylated and methylated DNA (0, 20, 40, 50, 60, 80, 100% methylated) were prepared and analyzed in parallel (Figure 4). As expected a decreased methylation rate results in a

decrease of the intensities of the detected fragments except the control fragment at $m/z = 1991$ which is formed independent of methylation and therefore has been utilized for signal normalization. In comparison to the methylation mixtures the methylation signals of the clinical samples show a different intensity ratio which is due to some adjacent CpG sites showing stronger comethylation than others as already indicated by the analysis of 10 clones from these samples. E.g., the presence of strong signals for fragments 6, 8, 9, 13 and 14 in tumor sample T1 (for nomenclature see Figure 1) indicates a high relative degree of comethylation around positions 122, 136, 138, 145, 152, 154, 258 and 269, respectively. These are exactly the CpG positions showing some degree of comethylation in most of the clones analyzed for this sample (Figure 3). The normalized relative intensities of the signals indicate a minimum of 50% methylation at these positions. In contrast the absence of or presence of weak signals for fragments 1, 3, 4 and 5 is explained by a low degree of comethylation at positions 32, 81, 96 and 104. Both observations correspond well to the clone data. A similar methylation pattern is observed for tumor sample T2. The lower intensity of the observed signals corresponds well to the lower number of clones showing comethylation for this sample (Figure 3). In case of the normal colon samples N1 and N2 the absence of comethylation is evidenced by both the direct method and clone analysis. In summary, the methylation information gained by direct analysis of clinical DNA samples compares well to the accumulated information obtained from clones of the same samples, albeit on a more qualitative level, i.e. the level of comethylation.

Discussion

DNA methylation markers have been discovered for a variety of clinical questions, most prominently cancer screening and monitoring. The validation of differential methylation patterns on large sample numbers represents a key step in the development of diagnostic test and requires both cost efficient and high throughput techniques. Herein we present a method for analyzing methylation patterns fulfilling these criteria. All of PCR, *in vitro* transcription and guanine-specific cleavage of the generated transcript are performed in one pot. It has been demonstrated that cloned bisulfite DNA can be analyzed directly without prior PCR. When analyzing plasmid

DNA directly the T7 primer tag can be omitted because many cloning systems include a promoter sequence within the vector. Plasmid DNA was linearized by restriction to avoid transcription of the entire plasmid. Although the information generated by analyzing the RNase T1 fragmentation pattern by MALDI-TOF is complementary to the methylation information contained in conventional bisulfite clone sequence traces, MALDI analysis does not require costly consumables like fluorescence labeled dNTPs.

In this study the methylation pattern of unmethylated DNA and methylated DNA has been analyzed. The methylation information contained in the respective RNase T1 fragment patterns is in complete accordance with the sequencing data generated in parallel. In addition, it has been shown that occasional incomplete bisulfite conversion is detected position-specifically by the appearance of specific shorter fragments appearing instead of an expected larger fragment. As the consequence no methylation information is lost as long as these fragments are detected faithfully, which can be accomplished using sequence-based tracking algorithms.

Since there is an upper and lower m/z limit in the MALDI-TOF detection there always will be fragments which cannot be detected. This also has been observed in the analysis of cloned bisulfite-treated colon DNA samples. However, the higher the mass range of the MALDI instrument, the higher will be the proportion of CpG positions, the methylation state of which can be unambiguously assigned. Even though, the presence of methylation over a stretch of neighboring CpG positions (comethylation) is most reliably detected even within a limited mass range as demonstrated by Figure 3. It is this comethylation of promoter regions which represents the most valuable information in the context of most clinical questions. The presence of comethylation of two or more neighboring CpG positions in the same molecule within clinical samples is readily detected by the direct method as demonstrated in Figure 4. The relative abundance of comethylation reveals itself in a semiquantitative manner by comparison of methylation signal intensities to those of DNA mixtures with defined methylation states and normalization against the signal generated from a methylation-independent control tag. The selectivity of the technique for detecting comethylation represents a significant advantage over the direct sequencing of bisulfite PCR products. Direct sequencing does not have the ability to distinguish specific methylation patterns from random methylation which usually does not contain clinically relevant information (17).

In summary, the combination of transcription, RNase T1 cleavage and fragment analysis by MALDI-TOF is suggested for methylation pattern analysis in bisulfite DNA either directly or after cloning.

Acknowledgement

We wish to thank the German Ministry of Education and Research (BMBF) for financial support of this study by a grant No. KW0110. We also wish to thank Daniela Friedrich, Institute for Zoo and Wildlife Research, for kindly performing the cloning experiments.

References

1. Boyes, J., Bird, A. (1992) Repression of genes by DNA methylation depends on CpG density and promoter strength: evidence for involvement of a methyl-CpG binding protein. *EMBO J.*, **11**, 327-333.
2. Riggs, A.D. (2002) X chromosome inactivation, differentiation, and DNA methylation revisited, with a tribute to Susumu Ohno. *Cytogenet. Genome Res.*, **99**, 17-24.
3. Recillas-Targa, F. (2002) DNA methylation, chromatin boundaries, and mechanisms of genomic imprinting. *Arch. Med. Res.*, **33**, 428-38.
4. Jones, P.A. (2002) DNA methylation in cancer. *Oncogene*, **21**, 5358-5360.
5. Michalowsky, L.A., Jones, P.A. (1989) DNA methylation and differentiation. *Environ. Health Perspect.*, **80**, 189-197.
6. Frommer, M., McDonald, L.E., Millar, D.S., Collis, C.M., Watt, F., Grigg, G.W., Molloy, P.L., Paul, C.L. (1992) A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc. Natl. Acad. Sci. U S A*, **89**, 1827-1831.
7. Nordhoff, E., Cramer, R., Karas, M., Hillenkamp, F., Kirpekar, F., Kristiansen, K., Roepstorff, P. (1993) Ion stability of nucleic acids in infrared matrix-assisted laser desorption/ionization mass spectrometry. *Nucleic Acids Res.*, **21**, 3347-3357.
8. Hartmer, R., Storm, N., Boecker, S., Rodi, C.P., Hillenkamp, F., Jurinke, C., van den Boom, D. (2003) RNase T1 mediated base-specific cleavage and MALDI-TOF MS for high-throughput comparative sequence analysis. *Nucleic Acids Res.*, **31**, e47.

- 9.Bocker,S. (2003) SNP and mutation discovery using base-specific cleavage and MALDI-TOF mass spectrometry. *Bioinformatics*, **19**, i44-53.
- 10.Krebs,S., Medugorac,I., Seichter,D., Forster,M. (2003) RNaseCut: a MALDI mass spectrometry-based method for SNP discovery. *Nucleic Acids Res.*, **31**, e37.
- 11.Seichter,D., Krebs,S., Forster M. (2004) Rapid and accurate characterisation of short tandem repeats by MALDI-TOF analysis of endonuclease cleaved RNA transcripts. *Nucleic Acids Res.*, **32**, e16.
- 12.Dean,F.B., Hosono,S., Fang,L., Wu,X., Faruqi,A.F., Bray-Ward,P., Sun,Z., Zong,Q., Du,Y., Du,J., Driscoll,M., Song,W., Kingsmore,S.F., Egholm,M., Lasken,R.S. (2002) Comprehensive human genome amplification using multiple displacement amplification. *Proc. Natl. Acad. Sci. U S A*, **99**, 5261-5266.
- 13.Olek,A., Oswald,J., Walter,J. (1996) A modified and improved method for bisulphite based cytosine methylation analysis. *Nucleic Acids Res.*, **24**, 5064-5066.
- 14.Toyooka,S., Toyooka,K.O., Harada,K., Miyajima,K., Makarla,P., Sathyanarayana,U.G., Yin,J., Sato,F., Shivapurkar,N., Meltzer,S.J., Gazdar,A.F. (2002) Aberrant methylation of the CDH13 (H-cadherin) promoter region in colorectal cancers and adenomas. *Cancer Res.*, **62**, 3382-3386.
- 15.Berkenkamp,S., Kirpekar,F., Hillenkamp,F. (1998) Infrared MALDI mass spectrometry of large nucleic acids. *Science*, **281**, 260-262.
- 16.Little,D.P., Thannhauser,T.W., McLafferty,F.W. (1995) Verification of 50- to 100-mer DNA and RNA sequences with high-resolution mass spectrometry. *Proc Natl Acad Sci U S A*, **92**, 2318-2322.
- 17.Song,J.Z., Stirzaker,C., Harrison,J., Melki,J.R., Clark,S.J. (2002) Hypermethylation trigger of the glutathione-S-transferase gene (GSTP1) in prostate cancer cells. *Oncogene*, **21**, 1048-1061.

Figures+ Legends

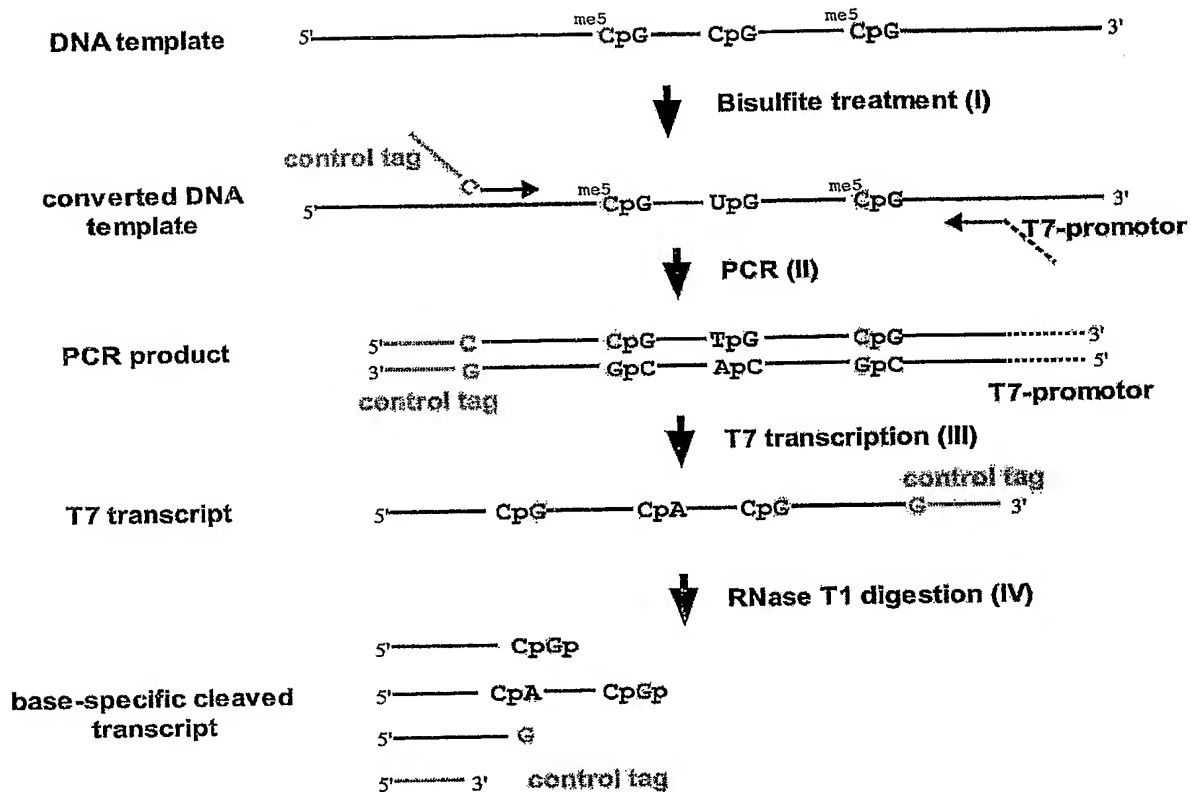


Figure 1: CpG methylation pattern analysis using RNase T1 cleavage and MALDI-TOF: (I) deamination of unmethylated cytosines to uracils by bisulfite treatment, (II) PCR amplification with the reverse primer containing a T7 promoter site and the forward primer carrying a control tag (yellow), (III) transcription to RNA containing G-sites at originally methylated cytosines (me⁵C), (IV) G-specific cleavage with RNase T1.

I.

```

1 3 5          32
A GGGAGCCCAAAUAAUACAACAACAUCACGAAACAUUAAUAAAAACUAAUACCAAAACC
B GGGAGCCCAAAUAAUACAACAACAUCACGAAACAUUAAUAAAAACUAAUACCAAAACC
C GGGAGCCCAAAUAAUACAACAACAUCACGAAACAUUAAUAAAAACUAAUACCAAAACC
D GGGAGCCCAAAUAAUACAACAACAUCACGAAACAUUAAUAAAAACUAAUACCAAAACC

81          96          104          122
A AAUACUUUACAAAACGAAUUCUUCUACGCUCCUUGUUUACAUAAACAAUACGAAUAA
B AAUACUUUACAAAACGAAUUCUUCUACGCUCCUUGUUUACAUAAACAAUACGAAUAA
C AAUACUUUACAAAACGAAUUCUUCUACGCUCCUUGUUUACAUAAACAAUACGAAUAA
D AAUACUUUACAAAACGAAUUCUUCUACGCUCCUUGUUUACAUAAACAAUACGAAUAA

136 138 145 152 154
A ACACCUCGCGAAAACGAAACCCCGGAAAUAACAUCACUUUACUUUUAAACUAAUAAAC
B ACACCUCACAAAACGAAACCCCGGAAAUAACAUCACUUUACUUUUAAACUAAUAAAC
C ACACCUCGCGAAAACGAAACCCCGGAAAUAACAUCACUUUACUUUUAAACUAAUAAAC
D ACACCUCGCGAAAACGAAACCCCGGAAAUAACAUCACUUUACUUUUAAACUAAUAAAC

210          242
A UCAACCUACAAAACGCUAAACAAUACCAUAUUCACUUUCCAAAACUAAUAAUACA
B UCAACCUACAAAACGCUAAACAAUACCAUAUUCACUUUCCAAAACUAAUAAUACA
C UCAACCUACAAAACGCUAAACAAUACCAUAUUCACUUUCCAAAACUAAUAAUACA
D UCAACCUACAAAACGCUAAACAAUACCAUAUUCACUUUCCAAAACUAAUAAUACA

258 266 269          292 298          308 310
A CGAAAACUAAACGACCAUCCAAACCUAAUACAA:AAAAA:AA
B CAAAACUAAACGACCAUCCAAACCUAAUACAA:AAAAA:AA
C CGAAAACUAAACGACCAUCCAAACCUAAUACAA:AAAAA:AA
D CGAAAACUAAACGACCAUCCAAACCUAAUACAA:AAAAA:AAUACAAUACUACUACU

```

II.

No.	Sequence	m/z	CpG position
P1	Gp	345	no CpG
P2	Gp	345	no CpG
P3	Gp	345	no CpG
P4	A ⁺ p	674	no CpG
1	CCCAAAUAAUACAACAACAACAUACGp	8619	32
2	AAAACAUAAUAAUAAACUAAUACCAAAACCAAAACUUUACAAAACGp	15724	32 and 81
3	AAUCCUUCUACGp	4719	81 and 96
4	CUCCUUGp	2484	96 and 104
5	UUUACAUACAAAACGp	5731	104 and 122
6	AAUAAACACCUUGp	4483	122 and 136
7	CGp	650	136 and 138
8	AAAAACGp	2296	138 and 145
9	AACCCCGp	2224	145 and 152
10	CGp	650	152 and 154
11	AAAUAAACAUCCGAUUUACUUUUUAAACUAAUAAACUACACCAAAACGp	17723	154 and 210
12	CUAAACAAUACCAACUAAUUCACUUUUCGAAAACUAAAUUACACGp	15253	210 and 258
13	AAAAACUACGp	3566	258 and 269
14	ACCAGUCCAAACCUAAUACAAAGp	7303	269
C1	AAAAACGp	1991	no CpG
C2	A	267	no CpG

III.

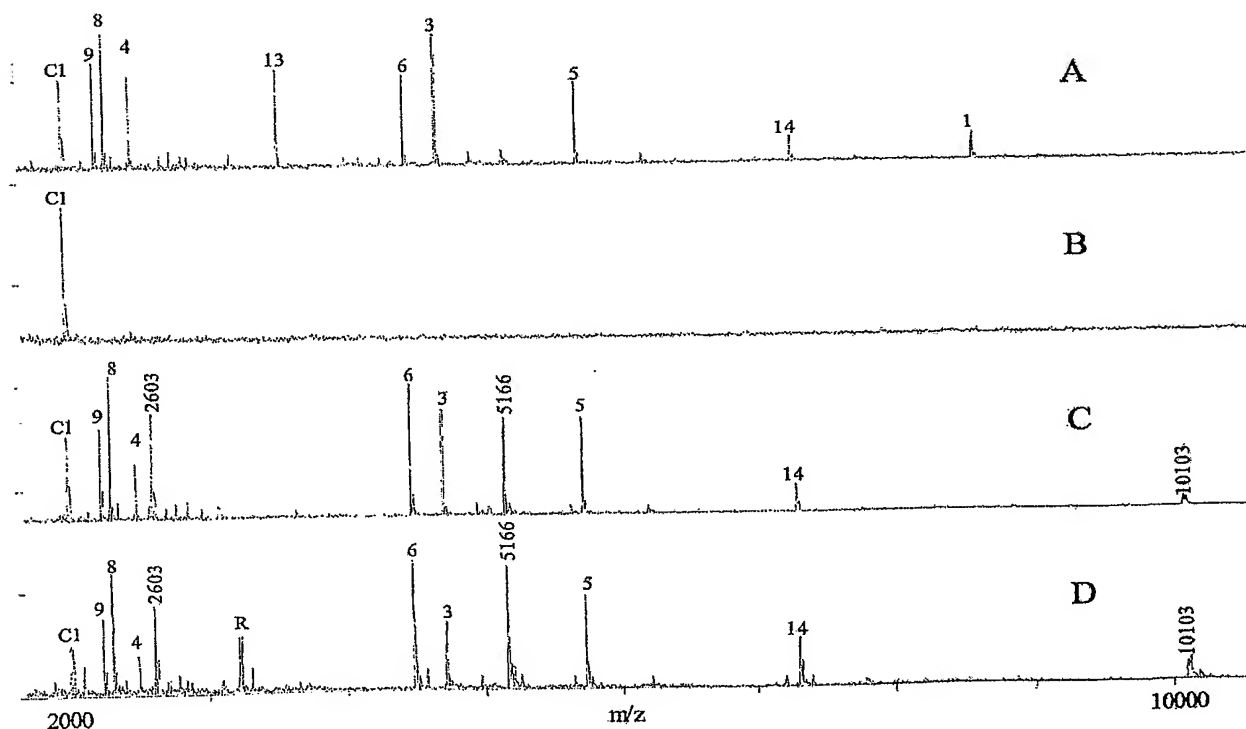


Figure 2: I.: Multiple alignment of the virtual sequences of the investigated T7 transcripts derived from the bisulfite sequence traces of: PCR product from methylated, bisulfite treated DNA (A), PCR product from unmethylated, bisulfite treated DNA (B), PCR product from plasmid pEPI2383 DNA (C) and plasmid

pEPI2383 without further PCR (D). Guanosines derived from originally methylated cytosines are highlighted in red, from nonconverted cytosines in blue and from the residual vector, T7 Promoter and control tag in yellow. Gene specific priming sites are marked in italic.

II.: Fragmentation (sequences, their related m/z values and positions of the represented CpG) of the T7 transcript A (in I.). Fragments from the T7 domain of the primer are tagged with P1 - P4, gene specific fragments are labelled with numbers (1-14) and fragments from the attached control tag are marked with C1-C2. Fragments size greater than 11000 Da and smaller than 1000 Da could not be detected. These are highlighted in italic. Mass accuracy for the mass range 2000 – 10000 m/z is ± 3 .

III.: MALDI TOF spectra of the base-specific cleaved T7 transcripts A, B, C and D (in I.)

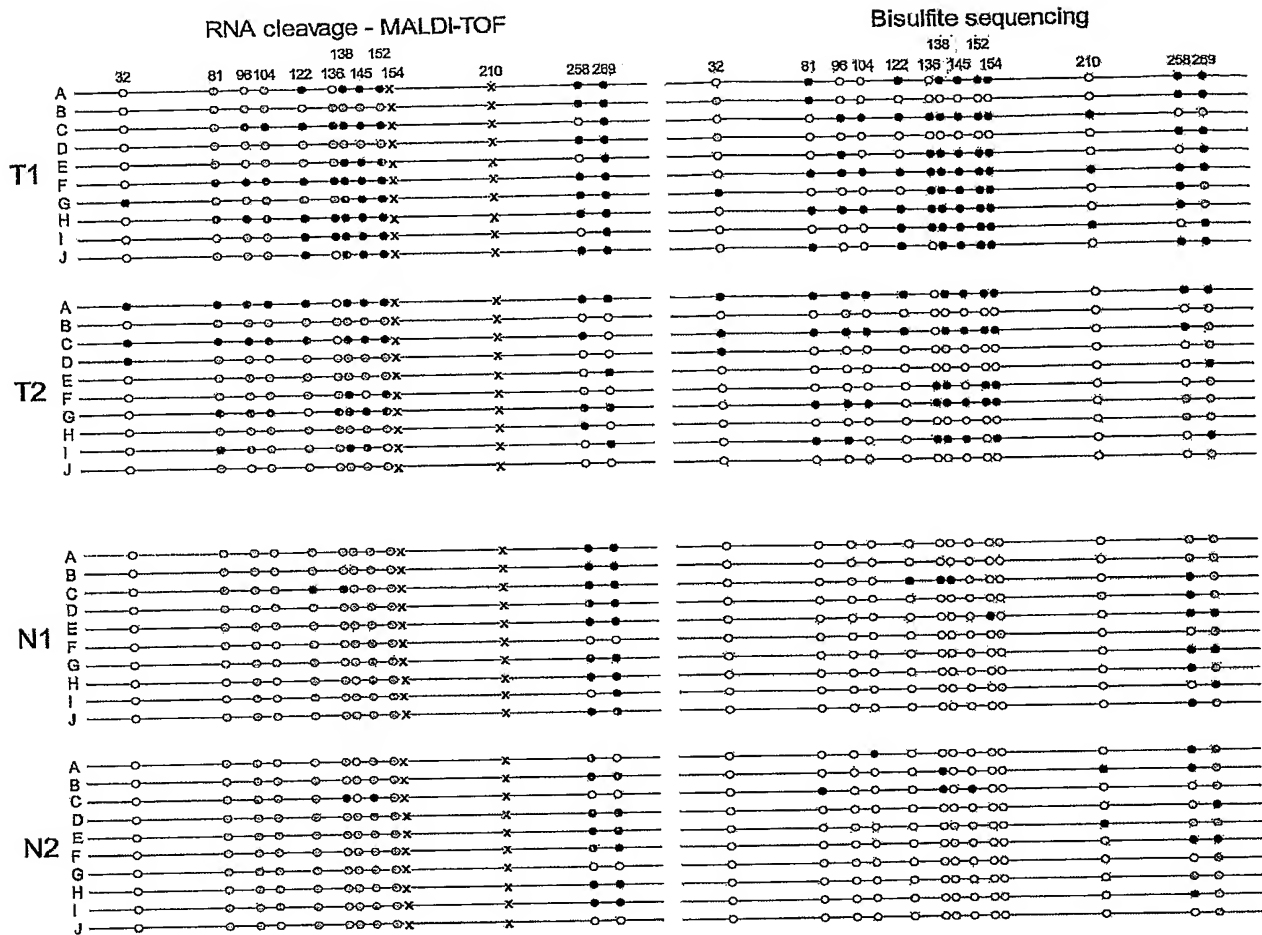


Figure 3: Analysis of CpG methylation in 10 clones (A-J) each derived from two bisulfite converted colon tumor DNA samples (T1, T2) and two normal colon DNA samples (N1, N2), resp., by RNA cleavage and MALDI-TOF (left) and sequencing (right). CpG methylated (black circle); CpG unmethylated (white circle); no fragment indicative of CpG methylation observed or ambiguous sequence (grey circle); CpG not accessible for analysis (cross).

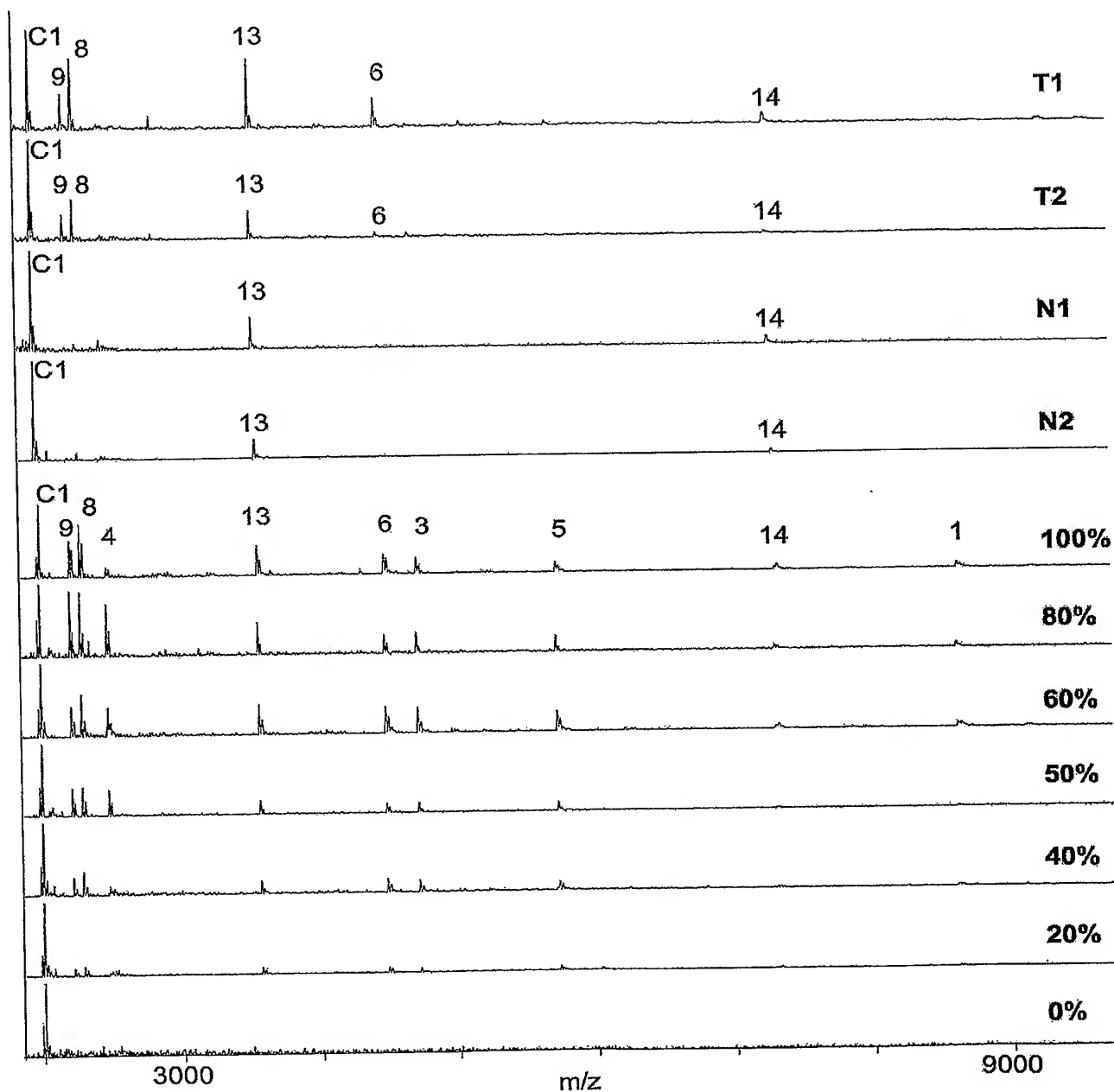


Figure 4: Direct analysis of two bisulfite converted colon tumor DNA samples (T1, T2) and two normal colon DNA samples (N1, N2), resp., by PCR, *in vitro* transcription, RNase T1 cleavage and MALDI-TOF in comparison to DNA mixtures with defined methylation states. (top: mass spectrum of samples T1, T2, N1, N2, resp.; bottom: spectra of DNA mixtures with different, defined methylation states – signal intensity of control fragment C1 set as max., fragment numbers given for 100%)

Patent Claim

1. Method for the analysis of cytosine methylation, characterised in that the following steps are conducted:
 - a) the DNA to be investigated is chemically treated in such a way that all of the unmethylated cytosine bases are converted to uracil or another base which is dissimilar to cytosine in terms of base pairing behaviour, while the 5-methylcytosine bases remain unchanged,
 - b) a promotor sequence is introduced into the DNA,
 - c) the DNA is transcribed into RNA, and
 - d) the RNA is analysed.